



## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANTS: C. Morrow et al.
- (ii) TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL NUCLEIC ACID AND METHODS OF MAKING AND USING SAME
- (iii) NUMBER OF SEQUENCES: 23
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: LAHIVE & COCKFIELD
  - (B) STREET: 28 STATE STREET
  - (C) CITY: BOSTON
  - (D) STATE: MASSACHUSETTS
  - (E) COUNTRY: USA
  - (F) ZIP: 02109
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: ASCII
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: 08/756,551
  - (B) FILING DATE: 08-JAN-2001
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 08/376,184
  - (B) FILING DATE: 17-AUG-1999
  - (C) CLASSIFICATION:
- (viii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 08/987,867
  - (B) FILING DATE: 09-DEC-1997
  - (C) CLASSIFICATION:
- (ix) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 08/389,459
  - (B) FILING DATE: 15-FEB-1995
- (x) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 08/087,009
  - (B) FILING DATE: 01-JUL-1993
- (xi) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Lauro, Peter C.
  - (B) REGISTRATION NUMBER: 32,360
  - (C) REFERENCE/DOCKET NUMBER: UAI-004CPDV2CN
- (xii) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (617) 227-7400
  - (B) TELEFAX: (617) 742-4214

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 14 base pairs
  - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TATTAGTAGA TCTG

14

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 14 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TACAGATGTA CTAA

14

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 845 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 20..845

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ACACAGCAAT CAGGTCAGC CAA AAT TAC CCT ATA GTG CAG AAC ATC CAG GGG 52  
 Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly  
 1 5 10

CAA ATG GTA CAT CAG GCC ATA TCA CCT AGA ACT TTA AAT GCA TGG GTA 100  
 Gln Met Val His Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val  
 15 20 25

AAA GTA GTA GAA GAG AAG GCT TTC AGC CCA GAA GTG ATA CCC ATG TTT 148  
 Lys Val Val Glu Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe  
 30 35 40

TCA GCA TTA TCA GAA GGA GCC ACC CCA CAA GAT TTA AAC ACC ATG CTA 196  
 Ser Ala Leu Ser Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu  
 45 50 55

AAC ACA GTG GGG GGA CAT CAA GCA GCC ATG CAA ATG TTA AAA GAG ACC 244  
 Asn Thr Val Gly Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr  
 60 65 70 75

ATC AAT GAG GAA GCT GCA GAA TGG GAT AGA GTG CAT CCA GTG CAT GCA 292

UAI-004CPDV2CN

Ile Asn Glu Glu Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala	80	85	90	
GGG CCT ATT GCA CCA GGC CAG ATG AGA GAA CCA AGG GGA AGT GAC ATA				340
Gly Pro Ile Ala Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile	95	100	105	
95	100	105		
GCA GGA ACT ACT AGT ACC CTT CAG GAA CAA ATA GGA TGG ATG ACA AAT				388
Ala Gly Thr Thr Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn	110	115	120	
110	115	120		
AAT CCA CCT ATC CCA GTA GGA GAA ATT TAT AAA AGA TGG ATA ATC CTG				436
Asn Pro Pro Ile Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu	125	130	135	
125	130	135		
GGA TTA AAT AAA ATA GTA AGA ATG TAT AGC CCT ACC AGC ATT CTG GAC				484
Gly Leu Asn Lys Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp	140	145	150	
140	145	150	155	
ATA AGA CAA GGA CCA AAG GAA CCC TTT AGA GAC TAT GTA GAC CGG TTC				532
Ile Arg Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe	160	165	170	
160	165	170		
TAT AAA ACT CTA AGA GCC GAG CAA GCT TCA CAG GAG GTA AAA AAT TGG				580
Tyr Lys Thr Leu Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp	175	180	185	
175	180	185		
ATG ACA GAA ACC TTG TTG GTC CAA AAT GCG AAC CCA GAT TGT AAG ACT				628
Met Thr Glu Thr Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr	190	195	200	
190	195	200		
ATT TTA AAA GCA TTG GGA CCA GCG GCT ACA CTA GAA GAA ATG ATG ACA				676
Ile Leu Lys Ala Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr	205	210	215	
205	210	215		
GCA TGT CAG GGA GTA GGA GGA CCC GGC CAT AAG GCA AGA GTT TTG GCT				724
Ala Cys Gln Gly Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala	220	225	230	
220	225	230	235	
GAA GCA ATG AGC CAA GTA ACA AAT TCA GCT ACC ATA ATG ATG CAG AGA				772
Glu Ala Met Ser Gln Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg	240	245	250	
240	245	250		
GGC AAT TTT AGG AAC CAA AGA AAG ATT GTT AAG TGT TTC AAT TGT GGC				820
Gly Asn Phe Arg Asn Gln Arg Lys Ile Val Lys Cys Phe Asn Cys Gly	255	260	265	
255	260	265		
AAA GAA GGG CAC ACA GCC AGA AAG T				845
Lys Glu Gly His Thr Ala Arg Lys	270	275		
270	275			

## (2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 275 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His Gln  
 1 5 10 15

Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu  
 20 25 30

Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu  
 35 40 45

Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly  
 50 55 60

His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala  
 65 70 75 80

Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro  
 85 90 95

Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser  
 100 105 110

Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro  
 115 120 125

Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile  
 130 135 140

Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly Pro  
 145 150 155 160

Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg  
 165 170 175

Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu  
 180 185 190

Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu  
 195 200 205

Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val  
 210 215 220

Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln  
 225 230 235 240

Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg Gly Asn Phe Arg Asn  
 245 250 255

Gln Arg Lys Ile Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Thr  
 260 265 270

Ala Arg Lys  
 275

## (2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 948 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 4..946

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AAC	CAA	TGG	CCA	TTG	ACA	GAA	GAA	AAA	ATA	AAA	GCA	TTA	GTA	GAA	ATT	48
Gln	Trp	Pro	Leu	Thr	Glu	Glu	Lys	Ile	Lys	Ala	Leu	Val	Glu	Ile		
1				5					10					15		
TGT	ACA	GAG	ATG	GAA	AAG	GAA	GGG	AAA	ATT	TCA	AAA	ATT	GGG	CCT	GAA	96
Cys	Thr	Glu	Met	Glu	Lys	Gly	Lys	Ile	Ser	Lys	Ile	Gly	Pro	Glu		
				20					25					30		
AAT	CCA	TAC	AAT	ACT	CCA	GTA	TTT	GCC	ATA	AAG	AAA	AAA	GAC	AGT	ACT	144
Asn	Pro	Tyr	Asn	Thr	Pro	Val	Phe	Ala	Ile	Lys	Lys	Lys	Asp	Ser	Thr	
				35				40						45		
AAA	TGG	AGA	AAA	TTA	GTA	GAT	TTC	AGA	GAA	CTT	AAT	AAG	AGA	ACT	CAA	192
Lys	Trp	Arg	Lys	Leu	Val	Asp	Phe	Arg	Glu	Leu	Asn	Lys	Arg	Thr	Gln	
				50				55						60		
GAC	TTC	TGG	GAA	GTT	CAA	TTA	GGA	ATA	CCA	CAT	CCC	GCA	GGG	TTA	AAA	240
Asp	Phe	Trp	Glu	Val	Gln	Leu	Gly	Ile	Pro	His	Pro	Ala	Gly	Leu	Lys	
				65				70						75		
AAG	AAA	AAA	TCA	GTA	ACA	GTA	CTG	GAT	GTG	GGT	GAT	GCA	TAT	TTT	TCA	288
Lys	Lys	Ser	Val	Thr	Val	Leu	Asp	Val	Gly	Asp	Ala	Tyr	Phe	Ser		
				80				85						95		
GTT	CCC	TTA	GAT	GAA	GAC	TTC	AGG	AAG	TAT	ACT	GCA	TTT	ACC	ATA	CCT	336
Val	Pro	Leu	Asp	Glu	Asp	Phe	Arg	Lys	Tyr	Thr	Ala	Phe	Thr	Ile	Pro	
									100					105		
AGT	ATA	AAC	AAT	GAG	ACA	CCA	GGG	ATT	AGA	TAT	CAG	TAC	AAT	GTG	CTT	384
Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	Val	Leu	
								115						125		
									120							

## UAI-004CPDV2CN

CCA CAG GGA TGG AAA GGA TCA CCA GCA ATA TTC CAA AGT AGC ATG ACA Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr 130 135 140	432
AAA ATC TTA GAG CCT TTT AGA AAA CAA AAT CCA GAC ATA GTT ATC TAT Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr 145 150 155	480
CAA TAC ATG GAT GAT TTG TAT GTA GGA TCT GAC TTA GAA ATA GGG CAG Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln 160 165 170 175	528
CAT AGA ACA AAA ATA GAG GAG CTG AGA CAA CAT CTG TTG AGG TGG GGA His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly 180 185 190	576
CTT ACC ACA CCA GAC AAA AAA CAT CAG AAA GAA CCT CCA TTC CTT TGG Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp 195 200 205	624
ATG GGT TAT GAA CTC CAT CCT GAT AAA TGG ACA GTA CAG CCT ATA GTG Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val 210 215 220	672
CTG CCA GAA AAA GAC AGC TGG ACT GTC AAT GAC ATA CAG AAG TTA GTG Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val 225 230 235	720
GGG AAA TTG AAT TGG GCA AGT CAG ATT TAC CCA GGG ATT AAA GTA AGG Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg 240 245 250 255	768
CAA TTA TGT AAA CTC CTT AGA GGA ACC AAA GCA CTA ACA GAA GTA ATA Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile 260 265 270	816
CCA CTA ACA GAA GAA GCA GAG CTA GAA CTG GCA GAA AAC AGA GAG ATT Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile 275 280 285	864
CTA AAA GAA CCA GTA CAT GGA GTG TAT TAT GAC CCA TCA AAA GAC TTA Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu 290 295 300	912
ATA GCA GAA ATA CAG AAG CAG GGG CAA GGC CTCGAG Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly 305 310	948

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 314 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys 1 5 10 15
Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn 20 25 30

Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys  
35 40 45

Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp  
50 55 60

Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys  
65 70 75 80

Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val  
85 90 95

Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser  
100 105 110

Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro  
115 120 125

Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys  
130 135 140

Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln  
145 150 155 160

Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His  
165 170 175

Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu  
180 185 190

Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met  
195 200 205

Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val Leu  
210 215 220

Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly  
225 230 235 240

Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln  
245 250 255

Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile Pro  
260 265 270

Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu  
275 280 285

Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile  
 290 295 300

Ala Glu Ile Gln Lys Gln Gly Gln Gly Leu  
 305 310

## (2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1568 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 7..1565

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGGGCC TGT CCA AAG GTA TCC TTT GAG CCA ATT CCC ATA CAT TAT TGT Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys 1 5 10	48
GCC CCG GCT GGT TTT GCG ATT CTA AAA TGT AAT AAT AAG ACG TTC AAT Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn 15 20 25 30	96
GGA ACA GGA CCA TGT ACA AAT GTC AGC ACA GTA CAA TGT ACA CAT GGA Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly 35 40 45	144
ATT AGG CCA GTA GTA TCA ACT CAA CTG CTG TTA AAT GGC AGT CTA GCA Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala 50 55 60	192
GAA GAA GAG GTA GTA ATT AGA TCT GTC AAT TTC ACG GAC AAT GCT AAA Glu Glu Glu Val Val Ile Arg Ser Val Asn Phe Thr Asp Asn Ala Lys 65 70 75	240
ACC ATA ATA GTA CAG CTG AAC ACA TCT GTA GAA ATT AAT TGT ACA AGA Thr Ile Ile Val Gln Leu Asn Thr Ser Val Glu Ile Asn Cys Thr Arg 80 85 90	288
CCC AAC AAC AAT ACA AGA AAA AGA ATC CGT ATC CAG AGA GGA CCA GGG Pro Asn Asn Asn Thr Arg Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly 95 100 105 110	336
AGA GCA TTT GTT ACA ATA GGA AAA ATA GGA AAT ATG AGA CAA GCA CAT Arg Ala Phe Val Thr Ile Gly Lys Ile Gly Asn Met Arg Gln Ala His 115 120 125	384

TGT AAC ATT AGT AGA GCA AAA TGG AAT AAC ACT TTA AAA CAG ATA GAT	432
Cys Asn Ile Ser Arg Ala Lys Trp Asn Asn Thr Leu Lys Gln Ile Asp	
130 135 140	
AGC AAA TTA AGA GAA CAA TTC GGA AAT AAT AAA ACA ATA ATC TTT AAG	480
Ser Lys Leu Arg Glu Gln Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys	
145 150 155	
CAA TCC TCA GGA GGG GAC CCA GAA ATT GTA ACG CAC AGT TTT AAT TGT	528
Gln Ser Ser Gly Gly Asp Pro Glu Ile Val Thr His Ser Phe Asn Cys	
160 165 170	
GGA GGG GAA TTT TTC TAC TGT AAT TCA ACA CAA CTG TTT AAT AGT ACT	576
Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr	
175 180 185 190	
TGG TTT AAT AGT ACT TGG AGT ACT GAA GGG TCA AAT AAC ACT GAA GGA	624
Trp Phe Asn Ser Thr Trp Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly	
195 200 205	
AGT GAC ACA ATC ACC CTC CCA TGC AGA ATA AAA CAA ATT ATA AAC ATG	672
Ser Asp Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met	
210 215 220	
TGG CAG AAA GTA GGA AAA GCA ATG TAT GCC CCT CCC ATC AGT GGA CAA	720
Trp Gln Lys Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln	
225 230 235	
ATT AGA TGT TCA TCA AAT ATT ACA GGG CTG CTA TTA ACA AGA GAT GGT	768
Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly	
240 245 250	
GGT AAT AGC AAC AAT GAG TCC GAG ATC TTC AGA CTT GGA GGA GGA GAT	816
Gly Asn Ser Asn Asn Glu Ser Gly Ile Phe Arg Leu Gly Gly Gly Asp	
255 260 265 270	
ATG AGG GAC AAT TGG AGA AGT GAA TTA TAT AAA TAT AAA GTA GTA AAA	864
Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys	
275 280 285	
ATT GAA CCA TTA GGA GTA GCA CCC ACC AAG GCA AAG AGA AGA GTG GTG	912
Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val	
290 295 300	
CAG AGA GAA AAA AGA GCA GTG GGA ATA GGA GCT TTG TTC CTT GGG TTC	960
Gln Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Leu Phe Leu Gly Phe	
305 310 315	
TTG GGA GCA GCA GGA AGC ACT ATG GGC GCA GCC TCA ATG ACG CTG ACG	1008
Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr	
320 325 330	
GTA CAG GCC AGA CAA TTA TTG TCT GGT ATA GTG CAG CAG CAG AAC AAT	1056
Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn	
335 340 345 350	

TTG CTG AGG GCT ATT GAG GCG CAA CAG CAT CTG TTG CAA CTC ACA GTC Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val 355 360 365	1104
TGG GGC ATC AAG CAG CTC CAA GCA AGA ATC CTA GCT GTG GAA AGA TAC Trp Gly Ile Lys Gln Leu Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr 370 375 380	1152
CTA AAG GAT CAA CAG CTC CTA GGG ATT TGG GGT TGC TCT GGA AAA CTC Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu 385 390 395	1200
ATT TGC ACC ACT GCT GTG CCT TGG AAT GCT AGT TGG AGT AAT AAA TCT Ile Cys Thr Thr Ala Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser 400 405 410	1248
CTG GAA CAG ATC TGG AAT CAC ACG ACC TGG ATG GAG TGG GAC AGA GAA Leu Glu Gln Ile Trp Asn His Thr Thr Trp Met Glu Trp Asp Arg Glu 415 420 425 430	1296
ATT AAC AAT TAC ACA AGC TTA ATA CAC TCC TTA ATT GAA GAA TCG CAA Ile Asn Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln 435 440 445	1344
AAC CAG CAA GAA AAG AAT GAA CAA GAA TTA TTG GAA TTA GAT AAA TGG Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp 450 455 460	1392
GCA AGT TTG TGG AAT TGG TTT AAC ATA ACA AAT TGG CTG TGG TAT ATA Ala Ser Leu Trp Asn Trp Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile 465 470 475	1440
AAA TTA TTC ATA ATG ATA GTA GGA GGC TTG GTA GGT TTA AGA ATA GTT Lys Leu Phe Ile Met Ile Val Gly Gly Leu Val Gly Leu Arg Ile Val 480 485 490	1488
TTT GCT GTA CTT TCT ATA GTG AAT AGA GTT AGG CAG GGA TAT TCA CCA Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro 495 500 505 510	1536
TTA TCG TTT CAG ACC CAC CTC CCA ATC TCGAG Leu Ser Phe Gln Thr His Leu Pro Ile 515	1568

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 519 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro  
 1 5 10 15

Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr  
 20 25 30

Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg  
 35 40 45

Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu  
 50 55 60

Glu Val Val Ile Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Thr Ile  
 65 70 75 80

Ile Val Gln Leu Asn Thr Ser Val Glu Ile Asn Cys Thr Arg Pro Asn  
 85 90 95

Asn Asn Thr Arg Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala  
 100 105 110

Phe Val Thr Ile Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn  
 115 120 125

Ile Ser Arg Ala Lys Trp Asn Asn Thr Leu Lys Gln Ile Asp Ser Lys  
 130 135 140

Leu Arg Glu Gln Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser  
 145 150 155 160

Ser Gly Gly Asp Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly  
 165 170 175

Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe  
 180 185 190

Asn Ser Thr Trp Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp  
 195 200 205

Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln  
 210 215 220

Lys Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg  
 225 230 235 240

Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn  
 245 250 255

Ser Asn Asn Glu Ser Glu Ile Phe Arg Leu Gly Gly Asp Met Arg  
 260 265 270

Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu  
 275 280 285

Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg  
 290 295 300  
 Glu Lys Arg Ala Val Gly Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly  
 305 310 315 320  
 Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln  
 325 330 335  
 Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu  
 340 345 350  
 Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly  
 355 360 365  
 Ile Lys Gln Leu Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys  
 370 375 380  
 Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys  
 385 390 395 400  
 Thr Thr Ala Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu  
 405 410 415  
 Gln Ile Trp Asn His Thr Thr Trp Met Glu Trp Asp Arg Glu Ile Asn  
 420 425 430  
 Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln  
 435 440 445  
 Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser  
 450 455 460  
 Leu Trp Asn Trp Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu  
 465 470 475 480  
 Phe Ile Met Ile Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala  
 485 490 495  
 Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser  
 500 505 510  
 Phe Gln Thr His Leu Pro Ile  
 515

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CACCCCTCTC CTACGTAACC AAGGATC

27

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GTACTGGTCA CCATATTGGT CAAC

24

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGAGAGAGAT GGGAGCTCGA GCGTC

25

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCCCCCCTAT ACGTATTGTG

20

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 41 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CCAGTGAATT CCTAATACGA CTCACTATAG GTTAAACAG C

41

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 48 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CTCTATCCTG AGCTCCATAT GTGTCGAGCA GTTTTTGGTT TAGCATTG

48

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Thr	Lys	Asp	Leu	Thr	Thr	Tyr	Gly
1				5			

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2220 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..2203

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGA	CCA	GCA	GAC	CAG	ACA	GTC	ACA	GCC	TTG	ACA	AAA	CGT	TCC	TGG	
Arg	Pro	Ala	Asp	Gln	Thr	Val	Thr	Ala	Ala	Leu	Thr	Lys	Arg	Ser	Trp
1				5				10				15			

48

AAC	TCA	AGC	ACT	TCT	CCA	CAG	AGG	AGG	ACA	GAG	CAG	ACA	GCA	GAG	ACC
Asn	Ser	Ser	Thr	Ser	Pro	Gln	Arg	Arg	Thr	Glu	Gln	Thr	Ala	Glu	Thr
20				25					30						

96

ATG	GAG	TCT	CCC	TCG	GCC	CCT	CCC	CAC	AGA	TGG	TGC	ATC	CCC	TGG	CAG
Met	Glu	Ser	Pro	Ser	Ala	Pro	Pro	His	Arg	Trp	Cys	Ile	Pro	Trp	Gln
35				40				45							

144

AGG	CTC	CTG	CTC	ACA	GCC	TCA	CTT	CTA	ACC	TTC	TGG	AAC	CCG	CCC	ACC
Arg	Leu	Leu	Leu	Thr	Ala	Ser	Leu	Leu	Thr	Phe	Trp	Asn	Pro	Pro	Thr
50				55				60							

192

ACT	GCC	AAG	CTC	ACT	ATT	GAA	TCC	ACG	CCG	TTC	AAT	GTC	GCA	GAG	GGG
Thr	Ala	Lys	Leu	Thr	Ile	Glu	Ser	Thr	Pro	Phe	Asn	Val	Ala	Glu	Gly
65			70				75					80			

240

AAG	GAG	GTG	CTT	CTA	CTT	GTC	CAC	AAT	CTG	CCC	CAG	CAT	CTT	TTT	GGC
Lys	Glu	Val	Leu	Leu	Val	His	Asn	Leu	Pro	Gln	His	Leu	Phe	Gly	
85			90			95									

288

TAC	AGC	TGG	TAC	AAA	GGT	GAA	AGA	GTG	GAT	GGC	AAC	CGT	CAA	ATT	ATA
Tyr	Ser	Trp	Tyr	Lys	Gly	Glu	Arg	Val	Asp	Gly	Asn	Arg	Gln	Ile	Ile
100					105							110			

336

GGA	TAT	GTA	ATA	GGA	ACT	CAA	CAA	GCT	ACC	CCA	GGG	CCC	GCA	TAC	AGT
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

384

Gly	Tyr	Val	Ile	Gly	Thr	Gln	Gln	Ala	Thr	Pro	Gly	Pro	Ala	Tyr	Ser	
115						120							125			
GGT	CGA	GAG	ATA	ATA	TAC	CCC	AAT	GCA	TCC	CTG	CTG	ATC	CAG	AAC	ATC	432
Gly	Arg	Glu	Ile	Ile	Tyr	Pro	Asn	Ala	Ser	Leu	Leu	Ile	Gln	Asn	Ile	
130						135							140			
ATC	CAG	AAT	GAC	ACA	GGA	TTC	TAC	ACC	CTA	CAC	GTC	ATA	AAG	TCA	GAT	480
Ile	Gln	Asn	Asn	Asp	Thr	Gly	Phe	Tyr	Thr	Leu	His	Val	Ile	Lys	Ser	
145						150							155			160
CTT	GTG	AAT	GAA	GAA	GCA	ACT	GGC	CAG	TTC	CGG	GTA	TAC	CCG	GAG	CTG	528
Leu	Val	Asn	Glu	Glu	Ala	Thr	Gly	Gln	Phe	Arg	Val	Tyr	Pro	Glu	Leu	
165						170							175			
CCC	AAG	CCC	TCC	ATC	TCC	AGC	AAC	AAC	TCC	AAA	CCC	GTG	GAG	GAC	AAG	576
Pro	Lys	Pro	Ser	Ile	Ser	Ser	Asn	Asn	Ser	Lys	Pro	Val	Glu	Asp	Lys	
180						185							190			
GAT	GCT	GTG	GCC	TTC	ACC	TGT	GAA	CCT	GAG	ACT	CAG	GAC	GCA	ACC	TAC	624
Asp	Ala	Val	Ala	Phe	Thr	Cys	Glu	Pro	Glu	Thr	Gln	Asp	Ala	Thr	Tyr	
195						200							205			
CTG	TGG	TGG	GTA	AAC	AAT	CAG	AGC	CTC	CCG	GTC	AGT	CCC	AGG	CTG	CAG	672
Leu	Trp	Trp	Val	Asn	Asn	Gln	Ser	Leu	Pro	Val	Ser	Pro	Arg	Leu	Gln	
210						215							220			
CTG	TCC	AAT	GGC	AAC	AGG	ACC	CTC	ACT	CTA	TTC	AAT	GTC	ACA	AGA	AAT	720
Leu	Ser	Asn	Gly	Asn	Arg	Thr	Leu	Thr	Leu	Phe	Asn	Val	Thr	Arg	Asn	
225						230							235			240
GAC	ACA	GCA	AGC	TAC	AAA	TGT	GAA	ACC	CAG	AAC	CCA	GTG	AGT	GCC	AGG	768
Asp	Thr	Ala	Ser	Tyr	Lys	Cys	Glu	Thr	Gln	Asn	Pro	Val	Ser	Ala	Arg	
245						250							255			
CGC	AGT	GAT	TCA	GTC	ATC	CTG	AAT	GTC	CTC	TAT	GGC	CCG	GAT	GCC	CCC	816
Arg	Ser	Asp	Ser	Val	Ile	Leu	Asn	Val	Leu	Tyr	Gly	Pro	Asp	Ala	Pro	
260						265							270			
ACC	ATT	TCC	CCT	CTA	AAC	ACA	TCT	TAC	AGA	TCA	GGG	GAA	AAT	CTG	AAC	864
Thr	Ile	Ser	Pro	Leu	Asn	Thr	Ser	Tyr	Arg	Ser	Gly	Glu	Asn	Leu	Asn	
275						280							285			
CTC	TCC	TGC	CAT	GCA	GCC	TCT	AAC	CCA	CCT	GCA	CAG	TAC	TCT	TGG	TTT	912
Leu	Ser	Cys	His	Ala	Ala	Ser	Asn	Pro	Pro	Ala	Gln	Tyr	Ser	Trp	Phe	
290						295							300			
GTC	AAT	GGG	ACT	TTC	CAG	CAA	TCC	ACC	CAA	GAG	CTC	TTT	ATC	CCC	AAC	960
Val	Asn	Gly	Thr	Phe	Gln	Gln	Ser	Thr	Gln	Glu	Leu	Phe	Ile	Pro	Asn	
305						310							315			320
ATC	ACT	GTG	AAT	AAT	AGT	GGA	TCC	TAT	ACG	TGC	CAA	GCC	CAT	AAC	TCA	1008
Ile	Thr	Val	Asn	Asn	Ser	Gly	Ser	Tyr	Thr	Cys	Gln	Ala	His	Asn	Ser	
325						330							335			
GAC	ACT	GGC	CTC	AAT	AGG	ACC	ACA	GTC	ACG	ACG	ATC	ACA	GTC	TAT	GCA	1056
Asp	Thr	Gly	Leu	Asn	Arg	Thr	Thr	Val	Thr	Thr	Ile	Thr	Val	Tyr	Ala	
340						345							350			
GAG	CCA	CCC	AAA	CCC	TTC	ATC	ACC	AGC	AAC	AAC	TCC	AAC	CCC	GTG	GAG	1104
Glu	Pro	Pro	Lys	Pro	Phe	Ile	Thr	Ser	Asn	Asn	Ser	Asn	Pro	Val	Glu	
355						360							365			
GAT	GAG	GAT	GCT	GTA	GCC	TTA	ACC	TGT	GAA	CCT	GAG	ATT	CAG	AAC	ACA	1152
Asp	Glu	Asp	Ala	Val	Ala	Leu	Thr	Cys	Glu	Pro	Glu	Ile	Gln	Asn	Thr	
370						375							380			
ACC	TAC	CTG	TGG	TGG	GTA	AAT	AAT	CAG	AGC	CTC	CCG	GTC	AGT	CCC	AGG	1200
Thr	Tyr	Leu	Trp	Trp	Val	Asn	Asn	Gln	Ser	Leu	Pro	Val	Ser	Pro	Arg	
385						390							395			400

CTG CAG CTG TCC AAT GAC AAC AGG ACC CTC ACT CTA CTC AGT GTC ACA Leu Gln Leu Ser Asn Asp Asn Arg Thr Leu Thr Leu Leu Ser Val Thr 405 410 415	1248
AGG AAT GAT GTA GGA CCC TAT GAG TGT GGA ATC CAG AAC GAA TTA AGT Arg Asn Asp Val Gly Pro Tyr Glu Cys Gly Ile Gln Asn Glu Leu Ser 420 425 430	1296
GTT GAC CAC AGC GAC CCA GTC ATC CTG AAT GTC CTC TAT GGC CCA GAC Val Asp His Ser Asp Pro Val Ile Leu Asn Val Leu Tyr Gly Pro Asp 435 440 445	1344
GAC CCC ACC ATT TCC CCC TCA TAC ACC TAT TAC CGT CCA GGG GTG AAC Asp Pro Thr Ile Ser Pro Ser Tyr Thr Tyr Tyr Arg Pro Gly Val Asn 450 455 460	1392
CTC AGC CTC TCC TGC CAT GCA GCC TCT AAC CCA CCT GCA CAG TAT TCT Leu Ser Leu Ser Cys His Ala Ala Ser Asn Pro Pro Ala Gln Tyr Ser 465 470 475 480	1440
TGG CTG ATT GAT GGG AAC ATC CAG CAA CAC ACA CAA GAG CTC TTT ATC Trp Leu Ile Asp Gly Asn Ile Gln Gln His Thr Gln Glu Leu Phe Ile 485 490 495	1488
TCC AAC ATC ACT GAG AAG AAC AGC GGA CTC TAT ACC TGC CAG GCC AAT Ser Asn Ile Thr Glu Lys Asn Ser Gly Leu Tyr Thr Cys Gln Ala Asn 500 505 510	1536
AAC TCA GCC AGT GGC CAC AGC AGG ACT ACA GTC AAG ACA ATC ACA GTC Asn Ser Ala Ser Gly His Ser Arg Thr Thr Val Lys Thr Ile Thr Val 515 520 525	1584
TCT GCG GAG CTG CCC AAG CCC TCC ATC TCC AGC AAC AAC TCC AAA CCC Ser Ala Glu Leu Pro Lys Pro Ser Ile Ser Ser Asn Asn Ser Lys Pro 530 535 540	1632
GTG GAG GAC AAG GAT GCT GTG GCC TTC ACC TGT GAA CCT GAG GCT CAG Val Glu Asp Lys Asp Ala Val Ala Phe Thr Cys Glu Pro Glu Ala Gln 545 550 555 560	1680
AAC ACA ACC TAC CTG TGG TGG GTA AAT GGT CAG AGC CTC CCA GTC AGT Asn Thr Thr Tyr Leu Trp Trp Val Asn Gly Gln Ser Leu Pro Val Ser 565 570 575	1728
CCC AGG CTG CAG CTG TCC AAT GGC AAC AGG ACC CTC ACT CTA TTC AAT Pro Arg Leu Gln Leu Ser Asn Gly Asn Arg Thr Leu Thr Leu Phe Asn 580 585 590	1776
GTC ACA AGA AAT GAC GCA AGA GCC TAT GTA TGT GGA ATC CAG AAC TCA Val Thr Arg Asn Asp Ala Arg Ala Tyr Val Cys Gly Ile Gln Asn Ser 595 600 605	1824
GTG AGT GCA AAC CGC AGT GAC CCA GTC ACC CTG GAT GTC CTC TAT GGG Val Ser Ala Asn Arg Ser Asp Pro Val Thr Leu Asp Val Leu Tyr Gly 610 615 620	1872
CCG GAC ACC CCC ATC ATT TCC CCC CCA GAC TCG TCT TAC CTT TCG GGA Pro Asp Thr Pro Ile Ile Ser Pro Pro Asp Ser Ser Tyr Leu Ser Gly 625 630 635 640	1920
GCG AAC CTC AAC CTC TCC TGC CAC TCG GCC TCT AAC CCA TCC CCG CAG Ala Asn Leu Asn Leu Ser Cys His Ser Ala Ser Asn Pro Ser Pro Gln 645 650 655	1968
TAT TCT TGG CGT ATC AAT GGG ATA CCG CAG CAA CAC ACA CAA GTT CTC Tyr Ser Trp Arg Ile Asn Gly Ile Pro Gln Gln His Thr Gln Val Leu 660 665 670	2016
TTT ATC GCC AAA ATC ACG CCA AAT AAC GGG ACC TAT GCC TGT TTT	2064

Phe Ile Ala Lys Ile Thr Pro Asn Asn Asn Gly Thr Tyr Ala Cys Phe	675	680	685	
GTC TCT AAC TTG GCT ACT GGC CGC AAT AAT TCC ATA GTC AAG AGC ATC	690	695	700	2112
Val Ser Asn Leu Ala Thr Gly Arg Asn Asn Ser Ile Val Lys Ser Ile				
ACA GTC TCT GCA TCT GGA ACT TCT CCT GGT CTC TCA GCT GGG GCC ACT	705	710	715	2160
Thr Val Ser Ala Ser Gly Thr Ser Pro Gly Leu Ser Ala Gly Ala Thr				
GTC GGC ATC ATG ATT GGA GTG CTG GTT GGG GTT GCT CTG ATA	725	730		2202
Val Gly Ile Met Ile Gly Val Leu Val Gly Val Ala Leu Ile				
TAGCAGCCCT GGTGTAGT				2220

## (2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 734 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Arg Pro Ala Asp Gln Thr Val Thr Ala Ala Leu Thr Lys Arg Ser Trp	1	5	10	15
Asn Ser Ser Thr Ser Pro Gln Arg Arg Thr Glu Gln Thr Ala Glu Thr	20	25		30
Met Glu Ser Pro Ser Ala Pro Pro His Arg Trp Cys Ile Pro Trp Gln	35	40	45	
Arg Leu Leu Leu Thr Ala Ser Leu Leu Thr Phe Trp Asn Pro Pro Thr	50	55	60	
Thr Ala Lys Leu Thr Ile Glu Ser Thr Pro Phe Asn Val Ala Glu Gly	65	70	75	80
Lys Glu Val Leu Leu Val His Asn Leu Pro Gln His Leu Phe Gly	85	90		95
Tyr Ser Trp Tyr Lys Gly Glu Arg Val Asp Gly Asn Arg Gln Ile Ile	100	105	110	
Gly Tyr Val Ile Gly Thr Gln Gln Ala Thr Pro Gly Pro Ala Tyr Ser	115	120	125	
Gly Arg Glu Ile Ile Tyr Pro Asn Ala Ser Leu Leu Ile Gln Asn Ile	130	135	140	
Ile Gln Asn Asp Thr Gly Phe Tyr Thr Leu His Val Ile Lys Ser Asp	145	150	155	160
Leu Val Asn Glu Glu Ala Thr Gly Gln Phe Arg Val Tyr Pro Glu Leu	165	170		175
Pro Lys Pro Ser Ile Ser Ser Asn Asn Ser Lys Pro Val Glu Asp Lys	180	185	190	
Asp Ala Val Ala Phe Thr Cys Glu Pro Glu Thr Gln Asp Ala Thr Tyr	195	200	205	
Leu Trp Trp Val Asn Asn Gln Ser Leu Pro Val Ser Pro Arg Leu Gln	210	215	220	

Leu Ser Asn Gly Asn Arg Thr Leu Thr Leu Phe Asn Val Thr Arg Asn  
 225 230 235 240  
 Asp Thr Ala Ser Tyr Lys Cys Glu Thr Gln Asn Pro Val Ser Ala Arg  
 245 250 255  
 Arg Ser Asp Ser Val Ile Leu Asn Val Leu Tyr Gly Pro Asp Ala Pro  
 260 265 270  
 Thr Ile Ser Pro Leu Asn Thr Ser Tyr Arg Ser Gly Glu Asn Leu Asn  
 275 280 285  
 Leu Ser Cys His Ala Ala Ser Asn Pro Pro Ala Gln Tyr Ser Trp Phe  
 290 295 300  
 Val Asn Gly Thr Phe Gln Gln Ser Thr Gln Glu Leu Phe Ile Pro Asn  
 305 310 315 320  
 Ile Thr Val Asn Asn Ser Gly Ser Tyr Thr Cys Gln Ala His Asn Ser  
 325 330 335  
 Asp Thr Gly Leu Asn Arg Thr Thr Val Thr Thr Ile Thr Val Tyr Ala  
 340 345 350  
 Glu Pro Pro Lys Pro Phe Ile Thr Ser Asn Asn Ser Asn Pro Val Glu  
 355 360 365  
 Asp Glu Asp Ala Val Ala Leu Thr Cys Glu Pro Glu Ile Gln Asn Thr  
 370 375 380  
 Thr Tyr Leu Trp Trp Val Asn Asn Gln Ser Leu Pro Val Ser Pro Arg  
 385 390 395 400  
 Leu Gln Leu Ser Asn Asp Asn Arg Thr Leu Thr Leu Leu Ser Val Thr  
 405 410 415  
 Arg Asn Asp Val Gly Pro Tyr Glu Cys Gly Ile Gln Asn Glu Leu Ser  
 420 425 430  
 Val Asp His Ser Asp Pro Val Ile Leu Asn Val Leu Tyr Gly Pro Asp  
 435 440 445  
 Asp Pro Thr Ile Ser Pro Ser Tyr Thr Tyr Tyr Arg Pro Gly Val Asn  
 450 455 460  
 Leu Ser Leu Ser Cys His Ala Ala Ser Asn Pro Pro Ala Gln Tyr Ser  
 465 470 475 480  
 Trp Leu Ile Asp Gly Asn Ile Gln Gln His Thr Gln Glu Leu Phe Ile  
 485 490 495  
 Ser Asn Ile Thr Glu Lys Asn Ser Gly Leu Tyr Thr Cys Gln Ala Asn  
 500 505 510  
 Asn Ser Ala Ser Gly His Ser Arg Thr Thr Val Lys Thr Ile Thr Val  
 515 520 525  
 Ser Ala Glu Leu Pro Lys Pro Ser Ile Ser Ser Asn Asn Ser Lys Pro  
 530 535 540  
 Val Glu Asp Lys Asp Ala Val Ala Phe Thr Cys Glu Pro Glu Ala Gln  
 545 550 555 560  
 Asn Thr Thr Tyr Leu Trp Trp Val Asn Gly Gln Ser Leu Pro Val Ser  
 565 570 575  
 Pro Arg Leu Gln Leu Ser Asn Gly Asn Arg Thr Leu Thr Leu Phe Asn  
 580 585 590

Val Thr Arg Asn Asp Ala Arg Ala Tyr Val Cys Gly Ile Gln Asn Ser  
 595 600 605  
 Val Ser Ala Asn Arg Ser Asp Pro Val Thr Leu Asp Val Leu Tyr Gly  
 610 615 620  
 Pro Asp Thr Pro Ile Ile Ser Pro Pro Asp Ser Ser Tyr Leu Ser Gly  
 625 630 635 640  
 Ala Asn Leu Asn Leu Ser Cys His Ser Ala Ser Asn Pro Ser Pro Gln  
 645 650 655  
 Tyr Ser Trp Arg Ile Asn Gly Ile Pro Gln Gln His Thr Gln Val Leu  
 660 665 670  
 Phe Ile Ala Lys Ile Thr Pro Asn Asn Asn Gly Thr Tyr Ala Cys Phe  
 675 680 685  
 Val Ser Asn Leu Ala Thr Gly Arg Asn Asn Ser Ile Val Lys Ser Ile  
 690 695 700  
 Thr Val Ser Ala Ser Gly Thr Ser Pro Gly Leu Ser Ala Gly Ala Thr  
 705 710 715 720  
 Val Gly Ile Met Ile Gly Val Leu Val Gly Val Ala Leu Ile  
 725 730

## (2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 41 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCAGTGAAATT CCTAATACGA CTACCTATAG GTTAAAACAG C

41

## (2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GATGAACCCT CGAGACCCAT TATG

24

## (2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:  
CCACCAAGTA CGTAACCACA TATGG 25

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 14 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:  
GTGAGGGACTG CTGG 14

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:  
CACCACTGCC CTCGAGAAGC TCACTATTG 29

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:  
CACCACTGCC CTCGAGAAGC TCACTATTG 29